

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-339019

(43)Date of publication of application : 06.12.1994

(51)Int.Cl.

H04N 1/40

G06K 9/20

H03M 7/30

H04N 1/41

(21)Application number : 05-126553

(71)Applicant : NIPPON TELEGR & TELEPH CORP  
<NTT>

(22)Date of filing : 28.05.1993

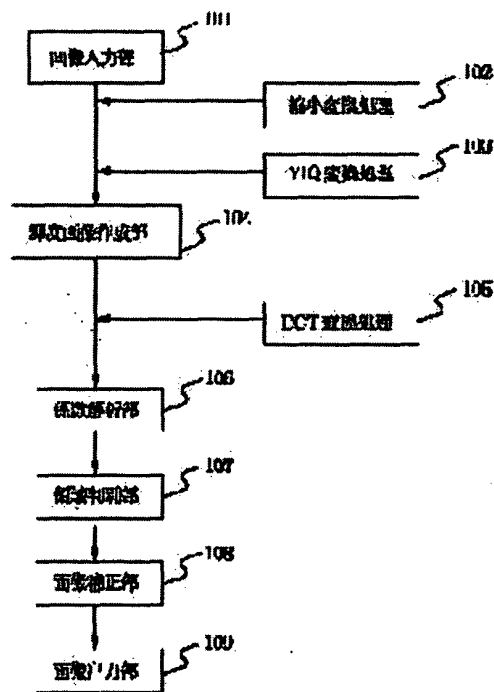
(72)Inventor : MIYAGAWA ISAO  
MATSUKI MAKOTO  
MIZUMACHI HAJIME

## (54) AREA SEPARATION SYSTEM FOR DOCUMENT PICTURE BY DISCRETE COSINE TRANSFORMATION

(57)Abstract:

**PURPOSE:** To separate a character area from a photographic area at a high speed by operating a picture reduction conversion, operating a discrete cosine transformation DCT processing, and discriminating a character from a photograph from a discrete cosine transformation coefficient obtained by the processing.

**CONSTITUTION:** The reduction conversion is operated to a full color picture by a reduction conversion processing 102, and Y components are searched by an YIQ transformation processing 103 by a luminance picture preparing part 104. The DCT processing 105 is operated to the Y components, and a coefficient value for discriminating a character from a photograph from a coefficient group is searched from among the 64 DCT coefficients by a counting analyzing part 106. That is, the absolute sum of the DCT coefficient and a high frequency coefficient is used as a characteristic coefficient according to the coefficient characteristic of the reduce picture. The DCT coefficient is largely affected by the influence of a background color, and set as a larger value than that of the photographic picture. Also, the diffusion of the high frequency components is large in the character part with more fine change, and the high frequency components are important at the time of reproduction. Thus, the separation of the character area from the photographic area including the background picture can be attained with a high quality by defining the characteristic coefficient as above mentioned.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

**\* NOTICES \***

**Japan Patent Office is not responsible for any damages caused by the use of this translation.**

- 1. This document has been translated by computer. So the translation may not reflect the original precisely.**
- 2. \*\*\*\* shows the word which can not be translated.**
- 3. In the drawings, any words are not translated.**

**CLAIMS**

**[Claim(s)]**

**[Claim 1] The field separation method of the document picture by the dispersed cosine conversion characterized by to distinguish a character and a photograph from the dispersed cosine transform coefficient which performed dispersed cosine transform processing after giving a picture contraction to the full color document picture acquired from the input unit to the printing document inputted into input units, such as an image scanner, in the field separation method which separates a character field and a photograph field, and was obtained by the aforementioned dispersed cosine transform processing.**

**[Translation done.]**

**\* NOTICES \***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

[Detailed Description of the Invention]

[0001]

[Industrial Application] this invention relates to the method which separates the field of a character and a photograph about the document with which a photograph etc. is intermingled to a character, a character, and the method that separates the field of a photograph, using dispersed cosine conversion (DCT being called hereafter) especially.

[0002]

[Description of the Prior Art] In the conventional field separation technology, the full color picture was changed into the binary picture, and the method which discriminates a character and a photograph field from the connection state of a pixel, and the method which asks for the texture of a picture in the state of the subject-copy image of a full color picture, and distinguishes the field of a character and a photograph existed. There are some which used FFT, a fractal dimension, wavelet transform, etc. for the latter typical thing.

[0003]

[Problem(s) to be Solved by the Invention] In the Prior art, in the picture incorporated from the image scanner, it is made binary, and with the end and the thing which distinguishes a character and a photograph in the state of connection of a pixel, while it has been full color, distinction of a character and a photograph is performed from the texture of a document picture. Although there is a field separation method adapting the Fourier transform, a fractal dimension, K-L conversion, wavelet transform, etc. as technique accepted by texture analysis, there is a trouble that all require time for conversion.

[0004] this invention aims at realizing highly precise and high-speed separation of a character and a photograph field by being made in view of the trouble which a Prior art which was mentioned above has, and using DCT.

[0005]

[Means for Solving the Problem] The field separation method of the document picture by dispersed cosine conversion of this invention performs dispersed cosine transform processing, after giving a picture contraction to the full color document picture acquired from the input unit to the printing document inputted into input units, such as an image scanner, in the field separation method which separates a character field and a photograph field, and it is characterized by to distinguish a character and a photograph from the dispersed cosine transform coefficient obtained by dispersed cosine transform processing.

[0006]

[Function] In order to fit DCT which attracts attention by the color picture coding method to texture analysis and to realize field separation recently, this invention First, adopt the reduction means of the given document picture and the coefficient by which a character and a photograph are characterized out of a DCT coefficient is perceived further. A character and a photograph field are distinguished for every block based on the performance index computed from the statistical property of the feature

coefficient in which it perceived, and it is characterized [ main ] by extracting a photograph field in the shape of a rectangle.

[0007] About the picture to which the picture contraction was carried out, a clear difference is accepted in each DCT coefficient property of a character picture and a photograph. In this invention, since DCT processing is performed about a reduction picture, field separation is performed good.

[0008]

[Example] Hereafter, the example of this invention is explained using drawing.

[0009] Drawing 1 is all the \*-block diagram of this invention, and drawing 2 and drawing 3 are drawings showing the detailed composition of this invention.

[0010] As shown in each drawing, the character inputted into the picture input section 101 (input unit 202) and the printing document 201 with which a photograph is intermingled are stored as a 24 bits (respectively every 8 bits) full color picture 203 of RGB.

[0011] Then, 1/[ 1/2 or ] 3 contraction is given by the contraction processing 102, the stored full color picture 203 is made into the reduction picture 204, and the Y component 205 is further called for by the YIQ transform processing 103 in the brightness picture creation section 104. In addition, although considered as the information which expresses a brightness component with this example using Y component as mentioned above, you may adopt the information which expresses a brightness component by making Y, Cr, Cb, and CIE(1976) L\*a\*b\* into a color space.

[0012] this invention -- this Y component 205 -- receiving -- the DCT processing 105 (a DCT block is a square whose one side is 8 pixels) -- giving -- counting -- in the analysis section 106, out of the DCT coefficient of 64 pieces, as shown in drawing 4 , the coefficient value for distinguishing a character and a photograph from a coefficient group is calculated In this example, the absolute sum of the DC coefficient and the RF coefficient which are shown in drawing 3 according to the coefficient property of a reduction picture was made into the feature coefficient. Especially DC coefficient is greatly influenced of a background color (especially white), and takes a bigger value than that of a photograph. Moreover, in the cover character section of change with the fine high frequency component shown with the slash in drawing 4 , distribution is large and it is an important portion in reproduction.

[0013] In this example, it is possible by having defined the feature coefficient as mentioned above to perform separation with the character field and photograph field containing the background image for high quality.

[0014] Next, distinction processing of a character and the photograph field is carried out based on the distinction performance index beforehand drawn analytically in the field distinction section 107.

[0015] Although the two-dimensional discriminant function (alignment) was applied in this invention, it is also possible to extend to n dimensions and to search for a performance index with combination with the coefficient considered to be an important coefficient, further. In addition, it asked for the discriminant function (alignment) used by this example according to the following procedures.

[0016]

[Equation 1]

$$dS^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = ax + by = M \dots\dots (1)$$

$$d = \begin{bmatrix} \{M(x_a) & M(y_a)\}/2 \\ \{M(x_b) & M(y_b)\}/2 \end{bmatrix}^T$$

$$S = \begin{bmatrix} \{V(x_a)+V(x_b)\}/2 & \{C(x_a,y_a)+C(x_b,y_b)\}/2 \\ \{C(x_a,y_a)+C(x_b,y_b)\}/2 & \{V(y_a)+V(y_b)\}/2 \end{bmatrix}$$

$$M = \{dS^{-1} \begin{bmatrix} M(x_a) \\ M(y_a) \end{bmatrix} + dS^{-1} \begin{bmatrix} M(x_b) \\ M(y_b) \end{bmatrix}\} / 2$$

但し、

$M(x_i)$  :  $x_i$  の平均値、 $V(x_i)$  :  $x_i$  の分散、 $C(x_i, x_j)$  :  $x_i$  と  $x_j$  の共分散

$x_a y_a$  : 写真画像クラスの AC 係数絶対和と DC 係数

$x_b y_b$  : 文字画像クラスの AC 係数絶対和と DC 係数

What is shown in a \*\* (1) formula is the performance index which distinguishes a character picture and a photograph. Therefore, field distinction of a character and a photograph asks for the DCT coefficient of each block, substitutes the feature coefficient value for x and y of a \*\* (1) formula, and if the value of left part is under M and it will become a photograph field and below M, it will distinguish from a character field. This distinction processing is performed about all blocks.

[0017] Next, the extraction method of a rectangle field is explained. Extraction processing shows detailed composition to drawing 3 including the field amendment after distinction (picture amendment).

[0018] As shown in drawing 3, when it is distinguished from a "photograph field" by the performance index, in the circumference of a photograph field, and the interior, the block distinguished from the character field exists and it is not rectangle-like depending on the case. Then, the interior of a field is amended by amendment processing using the adjoining field as shown in drawing 5 in the picture amendment section 108.

[0019] Next, in order to extract each field in the shape of a rectangle, labeling is performed and the maximum and the minimum value of a coordinate in every direction of each field as shown in drawing 6 are detected. More nearly finally than the picture output section 109, according to this coordinate, a field is extracted in rectangle, and it outputs.

[0020] In addition, it is for making easy structural coding performed after this, extracting a field in rectangle, as shown in drawing 6 may change according to the content of processing, when it is that from which future processings differ, and it is not especially important.

[0021]

[Effect of the Invention] As explained above, according to tree invention, it becomes possible to perform quality field separation of a character field (field containing the background region), and a photograph field at high speed to the document picture in which a character and a photograph are intermingled.

---

[Translation done.]

## AREA SEPARATION SYSTEM FOR DOCUMENT PICTURE BY DISCRETE COSINE TRANSFORMATION

Patent Number: JP6339019  
Publication date: 1994-12-06  
Inventor(s): MIYAGAWA ISAO; others: 02  
Applicant(s): NIPPON TELEGR & TELEPH CORP  
Requested Patent: ☐ JP6339019  
Application Number: JP19930126553 19930528  
Priority Number(s):  
IPC Classification: H04N1/40; G06K9/20; H03M7/30; H04N1/41  
EC Classification:  
Equivalents:

---

### Abstract

---

**PURPOSE:**To separate a character area from a photographic area at a high speed by operating a picture reduction conversion, operating a discrete cosine transformation DCT processing, and discriminating a character from a photograph from a discrete cosine transformation coefficient obtained by the processing.

**CONSTITUTION:**The reduction conversion is operated to a full color picture by a reduction conversion processing 102, and Y components are searched by an YIQ transformation processing 103 by a luminance picture preparing part 104. The DCT processing 105 is operated to the Y components, and a coefficient value for discriminating a character from a photograph from a coefficient group is searched from among the 64 DCT coefficients by a counting analyzing part 106. That is, the absolute sum of the DCT coefficient and a high frequency coefficient is used as a characteristic coefficient according to the coefficient characteristic of the reduce picture. The DCT coefficient is largely affected by the influence of a background color, and set as a larger value than that of the photographic picture. Also, the diffusion of the high frequency components is large in the character part with more fine change, and the high frequency components are important at the time of reproduction. Thus, the separation of the character area from the photographic area including the background picture can be attained with a high quality by defining the characteristic coefficient as above mentioned.

---

Data supplied from the esp@cenet database - I2

特開平6-339019

(43)公開日 平成6年(1994)12月6日

(51)IntCl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 1/40		F 9068-5C		
G 0 6 K 9/20	3 4 0	L		
H 0 3 M 7/30		A 8522-5J		
H 0 4 N 1/41		B 9070-5C		

審査請求 未請求 請求項の数1 O L (全 7 頁)

(21)出願番号 特願平5-126553

(22)出願日 平成5年(1993)5月28日

特許法第30条第1項適用申請有り 1993年3月15日、社団法人電子情報通信学会発行の「1993年電子情報通信学会春季大会講演論文集(7)」に発表

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72)発明者 宮川 勲

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72)発明者 松木 眞

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72)発明者 水町 肇

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

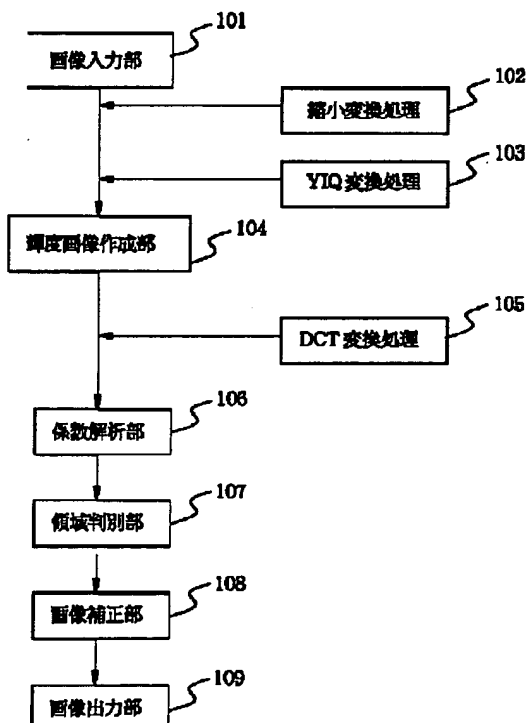
(74)代理人 弁理士 若林 忠

(54)【発明の名称】 離散的コサイン変換による文書画像の領域分離方式

(57)【要約】

【目的】 文字と写真領域の高精度かつ高速な分離を実現すること。

【構成】 イメージスキャナ等の入力装置に入力された印刷文書に対して、文字領域と写真領域を分離する領域分離方式において、入力装置から得られたフルカラー文書画像に対して画像縮小変換を施した後に離散的コサイン変換処理を行ない、前記離散的コサイン変換処理により得られた離散的コサイン変換係数から、文字と写真を判別することを特徴とする。





1

## 【特許請求の範囲】

【請求項1】 イメージスキャナ等の入力装置に入力された印刷文書に対して、文字領域と写真領域を分離する領域分離方式において、

入力装置から得られたフルカラー文書画像に対して画像縮小変換を施した後に離散的コサイン変換処理を行ない、

前記離散的コサイン変換処理により得られた離散的コサイン変換係数から、文字と写真を判別することを特徴とする離散的コサイン変換による文書画像の領域分離方式。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、文字と写真等が混在する文書から文字と写真の領域を分離する方式に関し、特に、離散的コサイン変換（以下、DCTと称する）を用いて文字と写真の領域を分離する方式に関する。

## 【0002】

【従来の技術】 従来の領域分離技術では、フルカラー画像を2値画像に変換して、画素の連結状態から文字と写真領域を識別する方式や、フルカラー画像の原画像状態で画像のテクスチャーを求め、文字と写真の領域を判別する方式が存在した。後者の代表的なものに、FFT、フラクタル次元、ウェーブレット変換等を利用したものがある。

## 【0003】

【発明が解決しようとする課題】 従来の技術では、イメージスキャナから取り込んだ画像を、一端、2値化し、画素の連結状態で文字と写真を判別するものと、フルカラーのまま、文書画像のテクスチャーから文字と写真の判別は実行されている。テクスチャー解析に通用される手法としては、フーリエ変換、フラクタル次元、K-L変換、ウェーブレット変換等を応用した領域分離方式があるが、いずれも変換に時間がかかるという問題点がある。

【0004】 本発明は上述したような従来の技術が有する問題点に鑑みてなされたものであってDCTを用いることにより、文字と写真領域の高精度かつ高速な分離を実現することを目的とする。

## 【0005】

【課題を解決するための手段】 本発明の離散的コサイン変換による文書画像の領域分離方式は、イメージスキャナ等の入力装置に入力された印刷文書に対して、文字領域と写真領域を分離する領域分離方式において、入力装置から得られたフルカラー文書画像に対して画像縮小変換を施した後に離散的コサイン変換処理を行ない、離散的コサイン変換処理により得られた離散的コサイン変換係数から、文字と写真を判別することを特徴とする。

## 【0006】

【作用】 本発明は、最近、カラー画像符号化方式で注目

2

されているDCTをテクスチャー解析に適応させ、領域分離を実現するために、まず、与えられた文書画像の縮小手段を採用し、さらに、DCT係数の中から文字と写真の特徴付ける係数に着眼して、着眼した特徴係数の統計的性質から算出した評価関数をもとに各ブロック毎に文字、写真領域を判別し、写真領域を矩形状に抽出することを最も主要な特徴とする。

【0007】 画像縮小変換が行われた画像については、文字画像と写真画像の各DCT係数特性に明確な差異が認められる。本発明では、縮小画像についてDCT処理が施されるので、領域分離が良好に行われる。

## 【0008】

【実施例】 以下、図を用いて本発明の実施例を説明する。

【0009】 図1は、本発明の全体的な構成図であり、図2、図3は、本発明の詳細な構成を示す図である。

【0010】 各図に示すように、画像入力部101（入力装置202）に入力された文字、写真が混在する印刷文書201は、RGBの24ビット（それぞれ8ビットずつ）のフルカラー画像203として格納される。

【0011】 この後、格納されたフルカラー画像203は、縮小変換処理102によって1/2、または、1/3縮小変換が施されて縮小画像204とされ、さらに、輝度画像作成部104でのYIQ変換処理103によってY成分205が求められる。なお、本実施例では上記のようにY成分を用いて輝度成分を表す情報としたが、この他に、Y、Cr、Cb、CIE(1976) L\*a\*b\*を色空間として輝度成分を表す情報を採用しても構わない。

【0012】 本発明は、このY成分205に対してDCT処理105（DCTブロックは一辺が8画素の正方形である）を施し、計数解析部106にて64個のDCT係数の中から、図4に示すように、係数群から文字と写真を判別するための係数値を求める。本実施例においては、縮小画像の係数特性に合わせて図3に示すDC係数と高周波係数との絶対和を特徴係数とした。DC係数は特に背景色の（とくに白）の影響を大きく受け、写真画像のそれよりも大きな値をとる。また、図4中の斜線で示した高周波成分は細かい変化の覆い文字部において分散が大きく、再生において重要な部分である。

【0013】 本実施例では、特徴係数を上記のように定義したことにより、背景画像を含んだ文字領域と写真領域との分離を高品質で行うことが可能となっている。

【0014】 次に、領域判別部107にて予め解析的に導出しておいた判別評価関数をもとに、文字と写真領域を判別処理する。

【0015】 本発明では、2次元の（線形）判別関数を適用したが、さらに、重要係数と考えられる係数との組み合わせによって、n次元に拡張し、それから評価関数を求めることも可能である。尚、本実施例で使用した

(線形) 判別関数は、以下の手順に従って求めた。

\*【数1】

【0016】

\*

$$dS^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = ax + by = M \dots\dots (1)$$

$$d = \begin{bmatrix} \{M(x_a) & M(y_a)\}/2 \\ \{M(x_b) & M(y_b)\}/2 \end{bmatrix}^T$$

$$S = \begin{bmatrix} \{V(x_a)+V(x_b)\}/2 & \{C(x_a,y_a)+C(x_b,y_b)\}/2 \\ \{C(x_a,y_a)+C(x_b,y_b)\}/2 & \{V(y_a)+V(y_b)\}/2 \end{bmatrix}$$

$$M = \{dS^{-1} \begin{bmatrix} M(x_a) \\ M(y_a) \end{bmatrix} + dS^{-1} \begin{bmatrix} M(x_b) \\ M(y_b) \end{bmatrix}\}/2$$

但し、

$M(x_i)$ :  $x_i$ の平均値、 $V(x_i)$ :  $x_i$ の分散、 $C(x_i, x_j)$ :  $x_i$ と $x_j$ の共分散

$x_a, y_a$ : 写真画像クラスのAC係数絶対和とDC係数

$x_b, y_b$ : 文字画像クラスのAC係数絶対和とDC係数

第(1)式に示すものが、文字画像と写真画像を判別する評価関数である。従って、文字と写真の領域判別は、各ブロックのDCT係数を求め、第(1)式の $x$ と $y$ に特徴係数値を代入して、左辺の値が $M$ 未満ならば写真領域、 $M$ 以下ならば文字領域と判別する。この判別処理は、全ブロックについて実行される。

【0017】次に、矩形領域の抽出方法について説明する。抽出処理は、判別後の領域補正(画像補正)を含むものであり、詳細な構成を図3に示す。

【0018】図3に示すように、評価関数により「写真領域」と判別された場合、写真領域の周辺、及び、内部において、文字領域と判別されたブロックが存在し、場合によっては矩形的でない。そこで、画像補正部108にて図5に示すような隣接する領域を用いた補正処理により、領域内部を補正する。

【0019】次に、個々の領域を矩形状に抽出するために、ラベリングを行ない、図6に示すような各領域の縦横の座標の最大値と最小値を検出する。最後に、この座標に従って矩形的に領域を抽出し、画像出力部109より出力する。

【0020】なお、図6に示したように領域を矩形的に抽出するのは、この後行われる構造的符号化を容易とするためであり、以後の処理が異なるものである場合には処理内容に応じて変更してもよく、特に重要なものではない。

【0021】

【発明の効果】以上説明したように、本発明によれば、文字と写真が混在する文書画像に対して、文字領域(背

景領域を含んだ領域)と写真領域の高品質な領域分離を高速に行うことが可能となる。

【図面の簡単な説明】

【図1】本発明の概略動作を示す図である。

【図2】本発明の判別処理までの構成図である。

【図3】本発明の判別処理以降の構成図である。

【図4】各ブロックで求めた64個のD印係数の中から、文字と写真を判別するために利用する重要係数を示した図であり、各格子はDCT係数 $F(m, n)$ に対応している。

【図5】画像補正に利用した穴埋め処理を示す図である。

【図6】補正された写真領域を矩形的に抽出する処理の説明図である。

【符号の説明】

101	画像入力部
102	縮小変換処理
103	YIQ変換処理
104	輝度画像作成部
105	DCT変換処理
106	係数解析部
107	領域判別部
108	画像補正部
109	画像出力部
201	文字、写真混在文書
202	入力装置
203	フルカラー画像
204	縮小画像

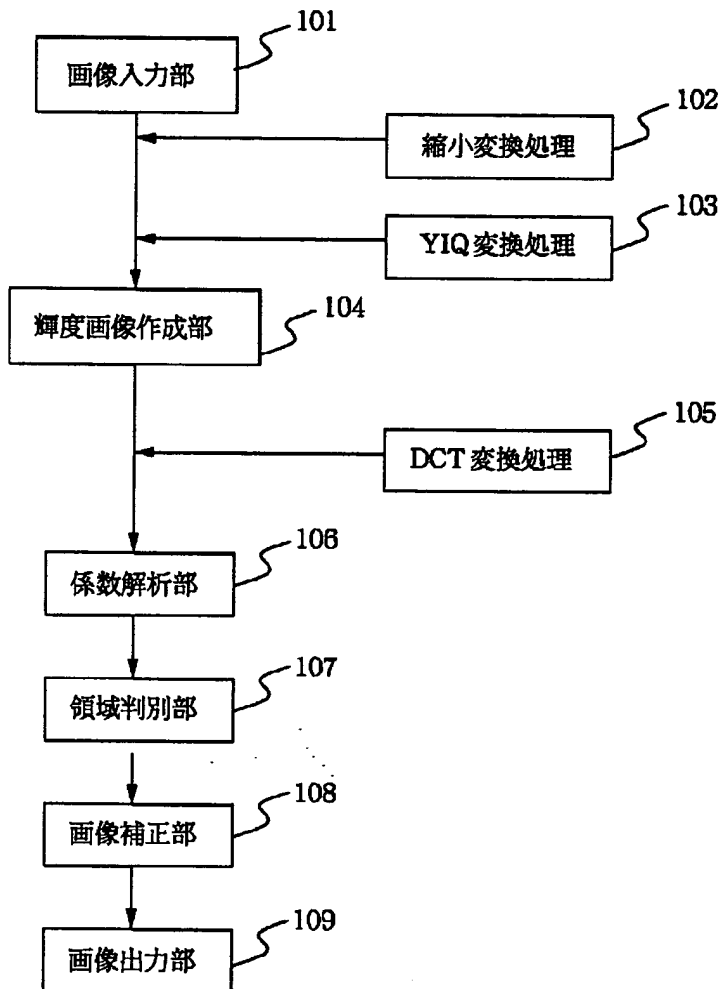
205 Y成分

5

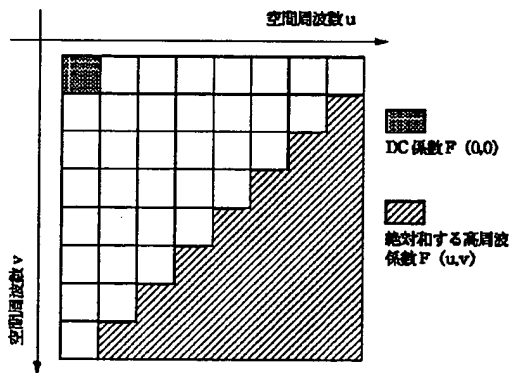
206 評価関数

6

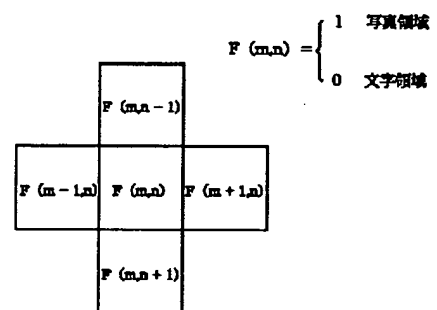
【図1】



【図4】

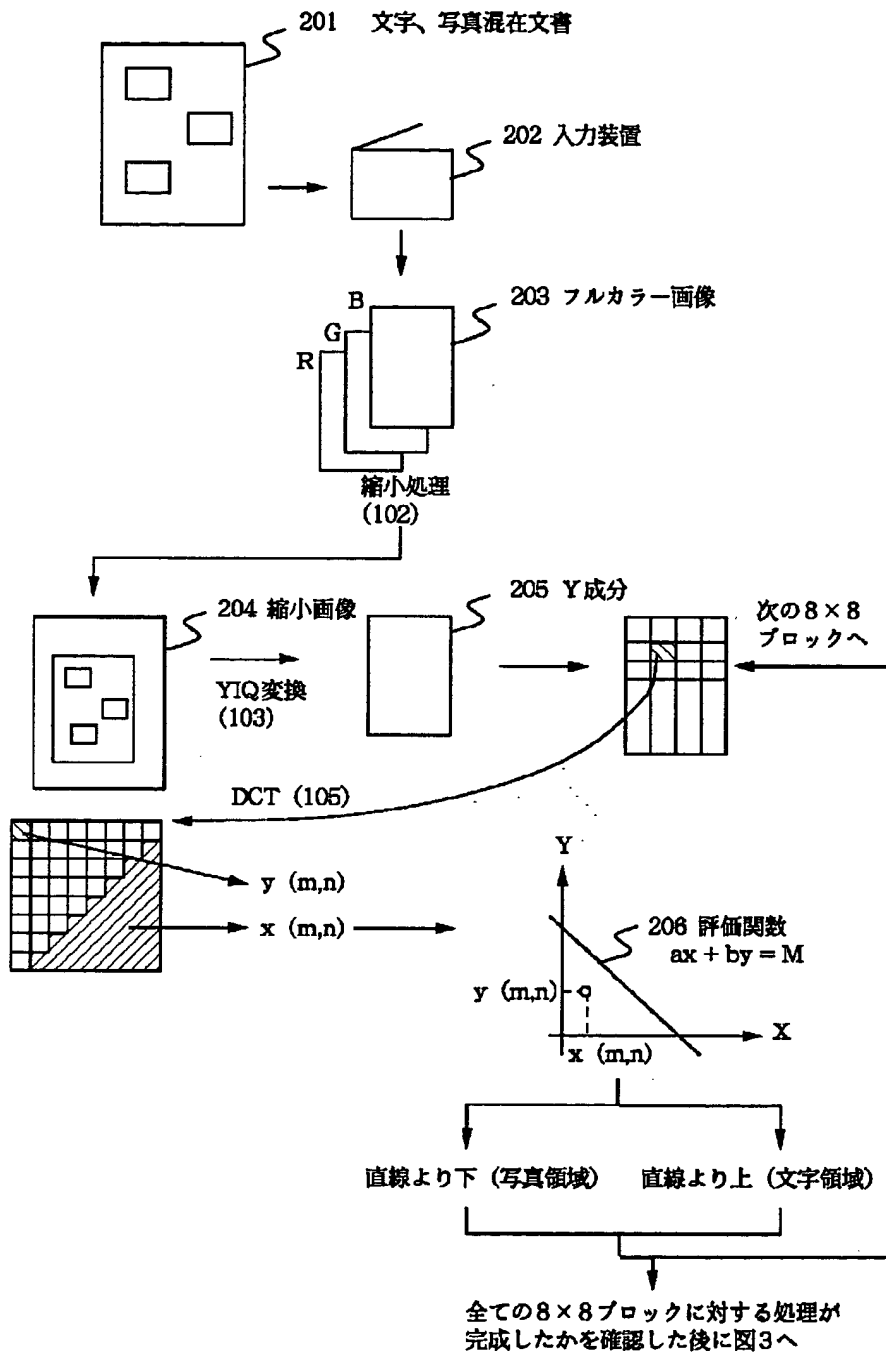


【図5】

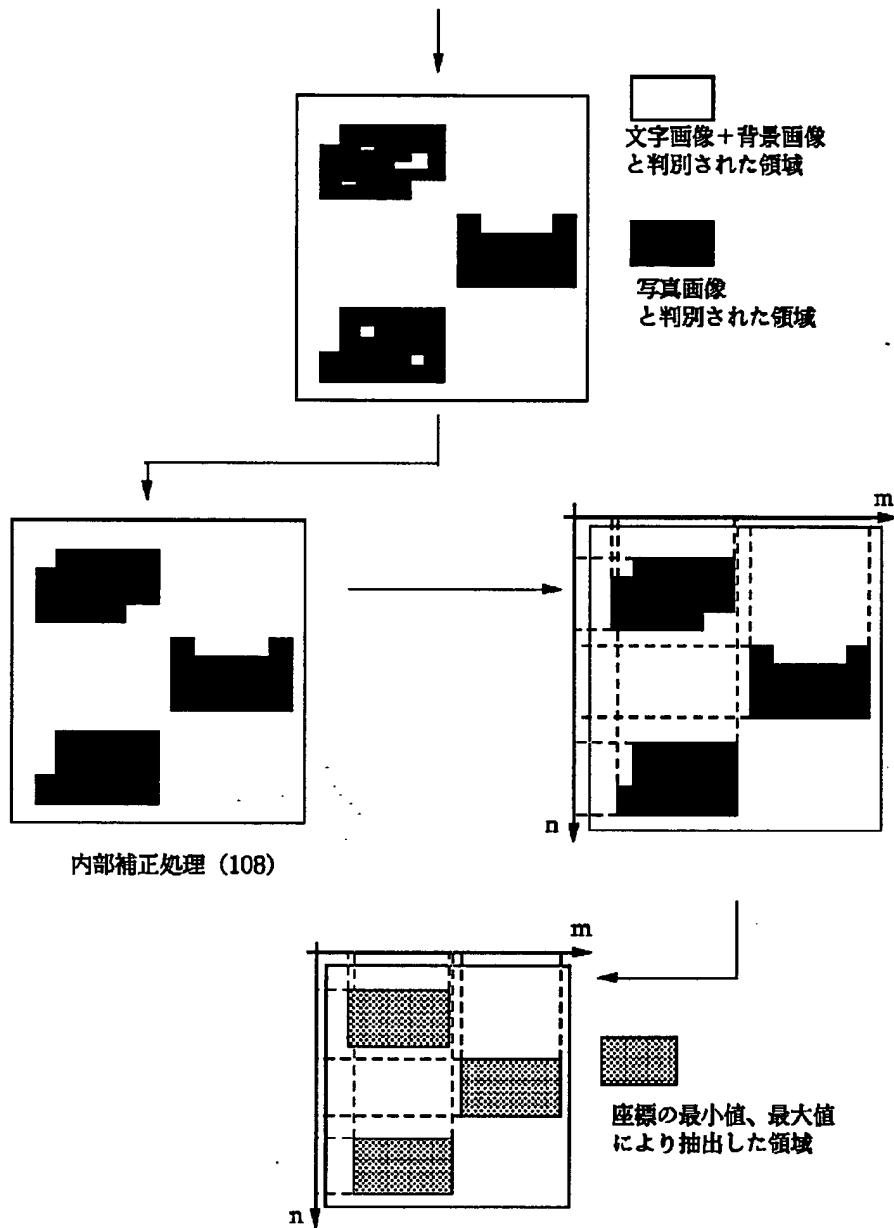


$$\begin{aligned} F(m,n-1) + F(m,n+1) + F(m-1,n) + F(m+1,n) \geq 2 & \quad F(m,n) = 1 \\ F(m,n-1) + F(m,n+1) + F(m-1,n) + F(m+1,n) < 2 & \quad F(m,n) = 0 \end{aligned}$$

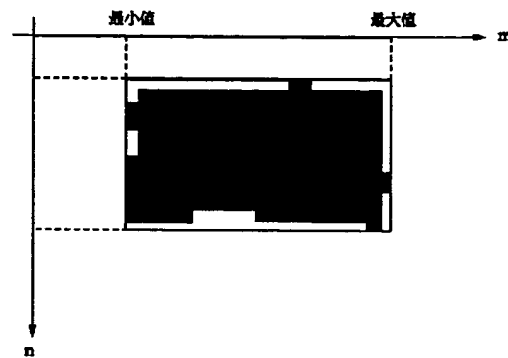
【図2】



【図3】



【図6】



■ 補正処理後の写真領域

□ 座標 (m,n) の最小値、最大値により抽出する領域